DOCUMENT RESUME

ED 341 706 TM 017 880

AUTHOR Blankenship, Mark H.; And Others
TITLE Evaluating the Economic Impact of a Valid Test: A Comparison of Supervisor and Incumbent Estimates of SDy.
PUB DATE Jun 91
NOTE 18p.; Paper presented at the Annual Meeting of the International Personnel Management Association Assessment Council (15th, Chicago, IL, June 23-27, 1991).
PUB TYPE Reports - Research/Technical (143) -- Speeches/Conference Papers (150)

EDRS PRICE MF01/PC01 Plus Postage.
DESCRIPTORS Comparative Analysis; *Cost Effectiveness; Economic Factors; Economic Impact; *Estimation (Mathematics); Evaluation Methods; *Evaluators; Government Employees; *Job Performance; Local Government; Personnel Selection; Questionnaires; *Supervisors; Test Validity
IDENTIFIERS California (San Diego County); Eligibility Technicians; Experts; *Job Value Factors; *Utility Analysis; Worth

ABSTRACT

To determine if individuals other than first-line supervisors could be used as expert judges in making estimates of job performance worth when conducting a utility analysis, this study compared the estimates of experienced job incumbents with those of supervisors. The job class investigated was that of eligibility technician, who performs varied technical work in determining clients' initial and continuing eligibility for one of several forms of public aid and is primarily found in the Department of Social Services in San Diego County (California). Participants were 120 eligibility supervisors and 203 eligibility technicians (incumbents). A modification of the global estimation technique of F. L. Schmidt and others was used to obtain estimates of job worth in dollars (SDy=job worth in dollars) at the 15th, 50th, and 85th percentiles of job performance. A questionnaire containing the modified Schmidt method was administered to the subjects. Results indicate that job incumbents provided estimates that were not substantially different from those of supervisors. The obtained estimates of worth were used in a utility analysis to compute costs and benefits associated with the eligibility technician selection process that was in use between December 1986 and January 1990; the process saved the county over one million dollars a year because it screened in high-performing individuals. One graph and 11 tables present study data, and a 35-item list of references is included. (SLD)

****
* Reproductions supplied by EDRS are the best that can be made *
* from the original document. *
****

U.S. DEPARTMENT OF EDUCATION
Office of Educational Research and Improvement
EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)
☑ This document has been reproduced as received from the person or organization originating it.
☐ Minor changes have been made to improve reproduction quality.
• Points of view or opinions stated in this document do not necessarily represent official OERI position or policy.

"PERMISSION TO REPRODUCE THIS MATERIAL HAS BEEN GRANTED BY
MARK H. BLANKENSHIP
TO THE EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)."

ED341706

Evaluating the Economic Impact of a Valid Test: A Comparison of Supervisor and Incumbent Estimates of SDy

Mark H. Blankenship
Department of Social Services
Management Services Bureau

Steven J. Cesare
Department of Human Resources
Recruitment and Assessment Division

Patrick W. Giannetto
Department of Human Resources
Recruitment and Assessment Division

Mark Z. Mandel
Department of Human Resources
Recruitment and Assessment Division

Running head: EVALUATING THE ECONOMIC IMPACT

This paper was part of the symposium entitled, A psychometric and fiscal assessment for selecting eligibility workers: Multiple agency results presented at the fifteenth annual conference of the International Personnel Management Association Assessment Council, Chicago, Illinois, June 1991.

TM017880

BEST COPY AVAILABLE

Abstract

This study was conducted to determine if individuals other than first-line supervisors could be used as expert judges in making estimates of job performance worth when conducting a utility analysis. Previous research is limited and the results have been mixed. Utility analysis can be used to weigh the costs and benefits associated with various human resource programs (e.g., selection and training) in economic terms. As such, it is a powerful tool for the industrial and organizational psychologist, permitting him/her to translate often esoteric statistical jargon into financial terms, thereby making it easier for management to understand the fiscal impact of human resource systems.

Recent methods for obtaining the information necessary to estimate utility have simplified the process. One such method, the global estimation procedure developed by Schmidt, Hunter, McKenzie and Muldrow (1979), has traditionally required that supervisors make estimate of worth based on various levels of performance. One problem with this method is that many small organizations typically do not have sufficient numbers of supervisors from which to obtain estimates. This is also the case in large organizations where the job class or job family is small.

This study attempted to determine if experienced job incumbents were capable of providing estimates of worth that were similar to those provided by supervisors. The job class investigated in this study was the Eligibility Technician, primarily found in the Department of Social Services. The participants were Eligibility Supervisors (n = 120), and Eligibility Technicians (n = 203). A modified Schmidt et al. global estimation method was used to obtain estimates of worth at the 15th, 50th and 85th percentiles of job performance. The results indicated that job incumbents provided estimates of job worth that were not statistically different from supervisorial estimates.

The obtained estimates of worth were used in a utility analysis to compute the costs and benefits associated with the Eligibility Technician selection process that was in use between December, 1986 and January, 1990. In addition to being statistically valid and unbiased, the selection process saved the County of San Diego over 1 million dollars per year in job performance. These savings were realized because the selection process screened in high performing individuals.

Evaluating the Economic Impact of a Valid Test: A Comparison of Supervisor and Incumbent Estimates of SDy

Research focusing on the application of utility analysis methods to estimate the costs and benefits of personnel selection and training programs has clearly demonstrated that valid programs can produce significant productivity and financial payoffs to organizations (Boudreau, 1989; Cascio, 1987; Cascio & Silbey, 1979; Cronshaw & Alexander, 1985; Hunter & Hunter, 1984; Landy, Farr, & Jacobs, 1982; Schmidt, Hunter, & Pearlman, 1982; Zeidner & Johnson, 1989). Despite these findings, industrial and organizational psychologists have been reluctant to apply theoretical procedures such as those developed by Brogden (1949) and Cronbach and Gleser (1965) for this purpose. According to Schmidt et al. (1982), the primary reason these procedures have not been employed is, "that one equation parameter--the *SD* [standard deviation] of job performance in dollars (SD_y)--had been considered extremely difficult to estimate" (p. 334). The assumption was that the SD_y estimate could only be obtained through costly and time consuming *cost accounting* procedures which, in the end, yielded only estimates of SD_y (Boudreau, 1989; Cascio, 1987; Cronbach & Gleser, 1965; Edwards, Frederick, & Burke, 1988; Roche, 1961; Schmidt et al., 1982).

The global estimation method for estimating SD_y, developed by Schmidt, Hunter, McKenzie, and Muldrow (1979) is one approach that has simplified SD_y estimation, and generated an increase in the number of utility analysis studies conducted applying Brogden-Cronbach-Gleser (B-C-G) formulas, which require SD_y estimates (Boudreau, 1989). Using the Schmidt et al. method, estimates of the dollar worth of job performance at the 15th, 50th, and 85th percentiles are obtained from expert judgements. Assuming that job performance is normally distributed, two estimates of SD_y are computed by taking the difference between the 85th and 50th percentile estimates (SD_{y1}) and the difference between the 50th and 15th percentile estimates (SD_{y2}) of worth. A third estimate of SD_y is found by computing the mean of SD_{y1} and SD_{y2}, this estimate is referred to as SD_{y3}. Furthermore, SD_{y3} has typically been the estimate of choice for computing utility when SD_{y1} and SD_{y2} are not statistically different.

Since the original study in 1979, several studies have tested the validity of SD_y estimates obtained by the Schmidt et al. method (Barrick, Barrett, Doverspike, Robison, & Grubs, 1990; Blankenship, Cesare, Giannetto, & Mandel, 1991; Bobko, Karren, & Parkington, 1983; Burke & Frederick, 1984, 1986; DeSimone, Alexander, & Cronshaw, 1986; Edwards, Frederick, & Burke, 1988; Greer & Cascio, 1987; Mayer, 1983; Reilly & Smither, 1985; Weekley, Frank, O'Connor, & Peters, 1985). Studies comparing the Schmidt et al. SD_y estimates to outside criteria (i.e., sales data, job performance data, cost accounting estimates, and estimates of overall worth), indicate that it may be a good approximation of outside criteria when those criteria are easily related to "dollar-values", but loses accuracy and becomes more variable as criteria become more difficult to translate into dollar terms (Reilly & Smither, 1985). Other studies, comparing Schmidt et al. SD_y estimates with SD_y estimates obtained from other methods have observed convergence when estimates of worth were primarily based on salary (Blankenship, et al., 1991; Edwards, et al., 1988; Reilly & Smither, 1985; Weekley, et al., 1985).

The Schmidt et al. method, as originally developed, relies on obtaining the requisite judgements of worth from supervisors of the target job class. Supervisors were chosen by Schmidt et al. because, "they have the best opportunities to observe actual performance and output differences between employees on a day-to-day basis" (p. 619). Thus, supervisors have a broad view of the organizational contributions made by incumbents of the target job class. Despite some psychometric criticisms, the use of supervisors as expert judges limits the Schmidt et al. method to large organizations, where ample numbers of subjects are more likely to be available (Bobko, et al., 1983). To date, few studies have compared Schmidt et al. estimates of SD_y obtained from multiple organizational levels (Matthews, Looper, & Engquist, 1990; Mayer, 1983) in order to ascertain if job incumbents provide estimates of SD_y that are comparable to supervisorial estimates.

Mayer (1983) had Tellers (job incumbents, n = 92), Head Tellers (supervisors, n = 49), and Branch/District Managers (managers, n = 51) from a large multibranch bank provide the necessary Schmidt et al. estimates of worth at the 15th, 50th, and 85th percentiles for the job class of Teller, during separate meetings held for each participant group. Results indicated that managers provided significantly higher estimates of worth than incumbents for the 50th and 85th percentiles, and for SD_{y1}, SD_{y2}, and SD_{y3}. Managers estimates of SD_{y2} were also significantly higher than corresponding supervisor estimates. By computing 95% confidence intervals around each of the three estimates of SD_{y3}, Mayer (1983) found that only the managers estimate overlapped a cost accounting estimate of worth.

Matthews, et al. (1990) had Task Scientists (job incumbents, n = 18) and their supervisors (n = 13) from a government human resources research laboratory provide Schmidt et al. estimates for the position of Task Scientist. The authors distributed a questionnaire to each participant and also made contact by phone to explain each question as posed by Schmidt et al., no guidance was provided. The results demonstrated that supervisors provided estimates of worth for the 50th and 85th percentiles, and for SD_{y1} that were significantly greater than corresponding estimates obtained from incumbents. An estimate for SD_{y3} was not computed because of SD_{y1} and SD_{y2} were significantly different for both supervisors and incumbents.

In sum, the paucity of research comparing supervisors' estimates of SD_y with those provided by job incumbents and/or managers provides a scant amount of evidence for choosing subject groups. In order to add to the existing research base, the purpose of this study was two-fold: to compare directly, SD_y estimates obtained from job incumbents and their immediate supervisors by using the Schmidt et al. (1979) method, and to then evaluate the economic utility of a selection procedure used by the County of San Diego to hire Eligibility Technicians between the years 1986 and 1990, based on the obtained SD_y estimates.

The hypotheses for this study were: (1) percentile estimates will not differ significantly by district office for supervisors or incumbents; (2) Schmidt et al. estimates of SD_{y1} and SD_{y2} will not be statistically different for supervisors or incumbents; and (3) there will be no significant differences between the SD_y estimates obtained from supervisors and incumbents.

Method

Definition of Relevant Job Class

Eligibility Technicians perform varied technical work in determining clients' initial and continuing eligibility for one of several forms of public aid: Aid to Families with Dependent Children, General Relief, Food Stamps, Refugee Assistance, and Medi-Cal. The combined programs amount to approximately 370 million dollars in annual aid within the County of San Diego.

Participants

One hundred and twenty-three Eligibility Supervisors (referred to as supervisors), and 230 Eligibility Technicians (referred to as incumbents) from the County of San Diego Department of Social Services participated in the study. Supervisors are one level above, and directly supervise Eligibility Technicians. Three of the supervisors and 8 of the incumbents failed to complete the questionnaire. Eleven of the incumbents failed to meet the minimum tenure as an Eligibility Technician (2 years) required to participate, and 8 more provided dollar estimates that were inconsistent with percentile magnitudes (Greer & Cascio, 1987; Weekley et al., 1985). As a result complete and usable data were obtained from 120 of the supervisors and 203 of the incumbents.

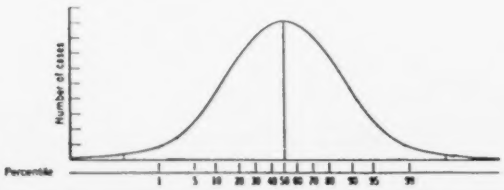
Experimental Design

As part of a larger study, a questionnaire containing a modified Schmidt et al. (1979) method for estimating SD_y was administered to the supervisors during regularly scheduled supervisors' meetings at each of the County's ten district offices (Blankenship, et al., 1991). A questionnaire containing the same modified Schmidt et al. (1979) method for estimating SD_y was then sent to 469 incumbents through the County mail.

SD_y Measures

The Schmidt et al. Method. The method developed by Schmidt et al. (1979) was used verbatim with two text changes: "Eligibility Technician" was substituted for "computer programmer" and "County of San Diego" was substituted for "my agency," and one addition. For each point estimate, participants were presented with an illustration of a normal distribution curve representing Eligibility Technician job performance, with a reference line drawn at the appropriate percentile, 50th, 85th, or 15th. A depiction of where the specified percentile was located in relation to a normal distribution was added to improve participants understanding of the questions (Bobko et al., 1983; Bobko, Karren, & Kerkar, 1987; Greer & Cascio, 1987). Instructions to the supervisors for estimating the job performance worth for the average performing Eligibility Technician were as follows (instructions for estimating the 85th and 15th percentiles were similar). The instructions were modified for the incumbents by replacing references to the participant as a supervisor with references to the participant as an incumbent.

> As part of this cost benefit study to improve selection methods for Eligibility Technicians, you will be asked to estimate the dollar value of products and services produced by Eligibility Technicians. These estimates are referred to as "dollar utility estimates."
>
> The dollar utility estimates we are asking you to make are critical in estimating the relative dollar value of different selection methods used by the County of San Diego. In answering these questions, you will have to make some very *difficult judgements*. We realize they are difficult and that they are judgements or estimates. You will have to ponder for some time before giving each estimate, and there is probably no way you can be absolutely certain your estimate is accurate when you do reach a decision. But keep in mind two things:
>
> (1) Your estimates will be averaged in with those of other supervisors of Eligibility Technicians. Thus, errors produced by too high and too low estimates will tend to be averaged out, providing more accurate final estimates.
>
> (2) The decisions that must be made about selection methods do not require that all estimates be accurate down to the last dollar. Substantially accurate estimates will lead to the same decisions as perfectly accurate estimates.
>
> Based on your experience with Eligibility Technicians, we would like for you to estimate the yearly value to the County of the products and services produced by the *average Eligibility Technician*. Consider the quality and quantity of output typical of the average Eligibility Technician and the value of this output. The graph below is a sample distribution of job performance for Eligibility Technicians. The line drawn at the 50th percentile represents the performance level for the *"average performing Eligibility Technician."*



low high

Eligibility Technician Job Performance

In placing an overall dollar value on this output, it may help to consider what the cost would be of having an outside firm provide these products and services.

Based on my experience, I estimate the value of the average Eligibility Technician to the County of San Diego at _____ dollars per year.

Three estimates of SD_y were computed for each treatment group (i.e., supervisors and incumbents). The first, SD_{y1} was computed as the mean difference between the 85th percentile and 50th percentile estimates. The second, SD_{y2} was computed as the mean difference between the 50th percentile and 15th percentile estimates. The third, SD_{y3} was computed as the average of SD_{y1} and SD_{y2} (Schmidt et al., 1979). The third estimate of SD_y was used as the overall estimate of SD_y in estimating utility.

Procedure

A letter was sent to all supervisors prior to the administration of the questionnaire. The letter explained that they would be asked to participate in a cost/benefit analysis project of the Eligibility Technician job class. While none of the supervisors reported any difficulty in understanding the Schmidt et al. questions, eight supervisors reported that they had difficulty in making the estimates. When asked whether or not a graphic presentation of the percentiles was useful, supervisors reported that the graphs helped clarify the percentile descriptions.

A similar letter was sent with a copy of the questionnaire to all incumbents (n = 469) who had been identified as having a minimum of 2 years experience as an Eligibility Technician with the County of San Diego. A two year minimum experience level was chosen for incumbent inclusion because incumbents are eligible for promotion to the supervisory level after that time. Questionnaires were returned by 230 of the incumbents for a 49 percent response rate.

Results

Treatment Group Demographics

Tenure statistics for each treatment group are shown in Table 1. Median tenure was included because of the large degree of variability about the supervisor group mean. Group data were also broken down by gender. The percentages of female supervisors and incumbents were 75.8 percent and 66.5 percent, respectively. The median number of subordinates for each supervisor was 8 (mean = 7.3, standard deviation [SD] = 2.9).

Table 1

Current Job Classification Tenure (in Years)

Group	Mean	Median	SD	Min	Max
Supervisors	3.0	1.8	4.0	0.1	18.4
Incumbents	4.5	4.0	1.8	2.0	12.0

Schmidt et al. (1979) Percentile Estimates

The mean value of worth for each percentile estimate of job performance was computed by district office. Supervisors estimates are shown in Table 2. As can be seen from the Table, supervisors felt that higher Eligibility Technician job performance was worth more to the organization than lower Eligibility Technician job performance.

Table 2

Supervisors' Mean Job Performance Percentile Estimates of Worth (in Dollars)

District Office	N	15th	50th	85th
Northeast	13	23,923	34,692	45,923
Lemon Grove	8	22,688	34,688	36,725
Logan Heights	11	21,718	31,636	39,445
South Bay	20	22,915	29,625	33,675
Kearny Mesa	20	20,220	29,850	41,430
Oceanside	11	19,955	27,091	34,318
Escondido	11	16,591	25,673	32,682
Southeast	10	21,100	30,000	37,500
El Cajon	12	21,375	28,250	35,533
Mission Valley	4	21,750	27,500	34,750
Total Group	120	21,255	29,962	37,535

A one-way multivariate analysis of variance (MANOVA) was conducted on the supervisors' percentile estimates. The independent factor was district office with 10 levels. The dependent variables were the 15th, 50th and 85th percentile estimates. The results were non-significant (Wilks' lambda likelihood ratio = .7582, $F(27,316.06) = 1.20^{1}$, $p < .235$) indicating that supervisor responses did not vary significantly by district office.

The mean value of worth for each percentile estimate of job performance was then computed by district office for incumbents. These values are shown in Table 3. As can be seen from Table 3, incumbents also felt that higher Eligibility Technician job performance was worth more to the organization than lower Eligibility Technician job performance.

[1] The F statistic associated with Wilks' lambda likelihood ratio statistic is Rao's approximation.

Table 3

Incumbents' Mean Job Performance Percentile Estimates of Worth (in Dollars)

District Office	N	15th	50th	85th
Northeast	22	17,532	27,900	36,586
Lemon Grove	14	22,000	28,500	34,964
Logan Heights	15	28,120	32,400	37,887
South Bay	19	24,884	33,184	42,495
Kearny Mesa	34	21,912	30,332	36,379
Oceanside	15	24,800	29,467	34,200
Escondido	28	21,700	27,943	34,618
Southeast	12	22,542	29,083	36,792
El Cajon	25	24,112	35,664	43,480
Mission Valley	15	22,267	26,933	32,267
Total Group[a]	199	22,651	30,195	37,088

[a] The total group sample size was less than 203 because 4 incumbents were from 3 offices other than the 10 district offices used for comparison. The group means were computed on all 203 incumbents.

Another one-way MANOVA was conducted on the incumbents' percentile estimates. The independent factor and dependent variables were the same as they were for the supervisors. The MANOVA results were significant (Wilks' lambda likelihood ratio = .7715, $F(27,546.78) = 1.88, p < .005$) indicating that incumbent responses varied significantly as a function of district office. To determine which dependent variable(s) differed significantly as a function of district office, a one-way univariate ANOVA was conducted on each dependent variable. Results indicated that only the 50th percentile estimate varied significantly. These results are shown in Table 4.

Table 4

One-Way Analysis of Variance Results for Incumbent Percentile Estimates

Dependent Variable	Source	SS	*df*	MS	*F*	*p*
15th Percentile	District Off.	1,294	9	144	1.89	.056
	Error	14,397	189	76		
50th Percentile	District Off.	1,469	9	163	1.99	.043
	Error	15,538	189	82		
85th Percentile	District Off.	2,316	9	257	1.04	.410
	Error	46,749	189	247		

Note. The data values were reduced by a factor of 1,000 to enhance readability.

Based on the above analyses, hypothesis one was accepted for supervisors' estimates of worth. Conversely, hypothesis one was not accepted for incumbents' estimates of worth since they varied significantly as a function of district office. Mean treatment group percentile estimates are shown in Table 5. There were no significant differences between supervisor and incumbent mean percentile estimates (see Table 6).

Table 5

Mean Treatment Group Job Performance Percentile Estimates of Worth (in Dollars)

Treatment Group	N	15th	50th	85th
Supervisors	120	21,255	29,962	37,535
Incumbents	203	22,651	30,195	37,088

Table 6

t-Tests of Supervisor vs. Incumbent Mean Job Performance Percentile Estimates

Percentile Estimate	*t*	*df*	*p*[b]
15th percentile	1.65[a]	309.64	.10
50th percentile	0.23	321.00	.82
85th percentile	0.29[a]	296.76	.78

[a] Separate variance estimates were used in computing the *t* statistic because of unequal treatment variances (Hayes, 1963).
[b] Two-tailed probability.

Schmidt et al. (1979) SD_y Estimates

Table 7 contains descriptive statistics for the Schmidt et al. SD_y estimates by treatment group. The means of SD_{y1} and SD_{y2} were compared to test hypothesis two. There were no significant differences for the supervisors ($t(119)$ .87, $p < .38$), or the incumbents ($t(202) = 1.39, p < .17$), suggesting that the distributions were symmetrical (Schmidt et al., 1979; Weekley et al., 1985). Based on the *t*-tests, hypothesis two was accepted. These tests could not confirm, however, that the distributions were normal (Bobko et al., 1983). Since there were no significant differences between SD_{y1} and SD_{y2} for the treatment groups, and in order to maintain consistency across studies, these estimates were averaged to obtain SD_{y3} for each group.

Table 7

Schmidt et al. SD_y Estimates (in Dollars)

Treatment Group	N	Mean	SD	SE
Supervisors				
SD_{y1}	120	7,832	7,845	716
SD_{y2}	120	8,448	6,660	608
SD_{y3}	120	8,140	6,166	563
Incumbents				
SD_{y1}	203	6,893	8,140	571
SD_{y2}	203	7,544	7,854	551
SD_{y3}	203	7,218	7,274	511

There were no significant differences between supervisor and incumbent SD_y estimates (see Table 8). Since SD_y estimates did not differ for supervisors and incumbents, hypothesis three was accepted.

Table 8

t-Tests of Supervisor vs. Incumbent SD_y Estimates

SD_y Estimate	*t*	*df*	*p*[b]
SD_{y1}	1.02	321.00	.31
SD_{y2}	1.10[a]	282.59	.27
SD_{y3}	1.21[a]	282.63	.23

[a] Separate variance estimates were used in computing the *t* statistic because of unequal treatment variances (Hayes, 1963).

[b] Two-tailed probability.

Pearson product-moment correlations were computed between the supervisors' demographic variables: tenure (in months), number of subordinates, and gender, with their corresponding percentile and SD_y estimates (see Table 9). There is no evidence from the correlations in Table 9 to support any significant degree of covariation between supervisors demographic characteristics and their percentile and SD_y estimates. Finally, similar correlations were computed for the incumbent treatment group (see Table 10). Results from these correlations are similar to those of the supervisors. Taken together, the correlations suggest that little covariation exists between the demographic variables used and percentile or SD_y estimates for any of the treatment groups. The findings reported here are similar to those reported by Schmidt et al. (1984) and Mayer (1983).

Table 9

Pearson Product-Moment Correlations Between Supervisors' Demographic Variables and Schmidt et al. SD_y Estimates

	15th	50th	85th	SD_{y1}	SD_{y2}	SD_{y3}
Length of Tenure	.007	-.086	-.114	-.128	-.066	-.117
Number of Subordinates	.010	.008	.065	.108	-.103	.013
Gender of Supervisor	.020	-.033	.030	.082	-.061	.020

Note. All correlations are non-significant $p > .05$, two-tailed ($n = 120$).

Table 10

Pearson Product-Moment Correlations Between Incumbents' Demographic Variables and Schmidt et al. SD_y Estimates

	15th	50th	85th	SD_{y1}	SD_{y2}	SD_{y3}
Length of Tenure	.012	-.037	-.083	-.117	-.057	-.096
Gender of Incumbent	-.092	-.034	.015	.066	.063	.071

Note. All correlations are non-significant $p > .05$, two-tailed ($n = 203$).

Computation of Selection Utility

In order to compute test utility it was necessary to determine: (a) the average tenure for an Eligibility Technician (in years); (b) the average number of persons hired to be Eligibility Technicians each year; (c) the validity of the selection process (corrected for range restriction and criterion unreliability); (d) the standard deviation of job performance in dollars (SD_y); (e) the height of the unit normal distribution [N(0,1)] at the test point of cut; (f) the selection ratio; and (g) the cost of testing an applicant.

The average tenure for an Eligibility Technician was computed by calculating the time in job class (in months) for every person who held an appointment as an Eligibility Technician and then left that job class. Based on a sample of 1,167 individuals, average tenure was estimated to be 28 months (2.33 years). The average number of Eligibility Technicians hired each year is approximately 200. This figure was based

on hiring numbers between 1986 and 1990. Test validity, computed by means of a predictive validation study, was found to be .3455 (Cesare, Blankenship, Mandel, & Giannetto, 1991). The standard deviation of job performance in dollars used was estimated in this study (see Table 7). The height of the unit normal curve was taken from the Naylor-Shine tables (Cascio, 1987), and was equal to .3034. The selection ratio was calculated to be .23 (938 total hires divided by 4,061 total applicants). Finally, the cost of selecting an applicant was estimated at $11.70.

The B-C-G equation used in computing test utility was:

$$\Delta U = tN_s (r_1 - r_2) SD_y (\lambda_i/\phi_i) - N_s (C_1 - C_2)/\phi_i$$

where ΔU is the marginal utility of the selection process, t is the average tenure for an Eligibility Technician, N_s is the number of Eligibility Technicians hired each year, r_1 and r_2 are the validity coefficients associated with the selection methods being compared, SD_y is as previously defined, λ_i is the height of the unit normal curve, ϕ_i is the selection ratio, and C_1 and C_2 are the cost of testing a single applicant for each of the selection methods associated with r_1 and r_2, respectively. The marginal utility estimated from the above formula is for one years use of the selection process. The savings would be realized over average tenure.

By substitution, marginal utility for the Eligibility Technician selection process is:

$$\Delta U = [2.33 \times 200 \times (.3455) \times SD_y \times (.3034 / .23)] - [200 \times ($11.70) / .23]$$

Because only one selection process is being considered, r_2 and C_2 drop out of the equation. Substituting the three estimates of SD_y, shown in Table 7, the total computed utility for the Eligibility Technician selection process ranges between 1.5 million dollars and 1.7 million dollars (see Table 11) in job performance (productivity) saved over a random selection process. These estimates translate into job performance savings of over 3,600 dollars per Eligibility Technician per year.

Table 11

Selection Process Utility (in Dollars)

Treatment Group	Years of Test Use	
	One Year	Four Years
Supervisors	1,719,771	6,879,084
Incumbents	1,523,824	6,095,296

Discussion

As stated in the introduction, the purpose of the present study had two objectives. The first objective was to compare SD_y estimates obtained from experienced job incumbents and supervisors using a modified Schmidt et al. (1979) procedure. A comparison of SD_y estimates across organizational levels was made to determine if employees other than supervisors could be used as expert judges when making SD_y estimates. Available research has demonstrated that significant differences in SD_y estimates are observed

across organizational structure (Matthews, et al., 1990; Mayer, 1983). The second objective was to estimate the utility of a selection procedure used to hire Eligibility Technicians in the County of San Diego between January 1986 and December 1989 based on the obtained estimates of SD_y.

Percentile Estimates

Excluding incumbents' 50th percentile estimates, the results indicated that percentile estimates were not statistically different across district offices for supervisors and incumbents. If these findings can be generalized, it may not be necessary to stratify the subject samples by district offices or geographic locations. This generalization would benefit large organizations where offices are separated by large distances. In addition to within group consistency, there were no significant mean differences between supervisor and incumbent estimates (Table 6). Similar to these findings, Mayer (1983) reported that supervisors and incumbents provided similar percentile estimates.

SD_y Estimates

A primary assumption of the Schmidt et al. method is that job performance is normally distributed (Schmidt et al., 1979; Vance & Colella, 1990). While this assumption could not be tested directly, results indicated that the distributions for supervisors and incumbents were at least symmetrical (Weekley, et al., 1985). The addition of a graphic presentation of a normal distribution with appropriate percentile location may have helped promote correct percentile magnitude estimation (Bobko et al., 1983) and distribution symmetry. Neither percentile estimates of worth nor SD_y estimates were related to either supervisory or incumbent demographic characteristics (e.g., length of tenure, number of subordinates [supervisors only], and gender), similar to the results reported by Schmidt, Mack, and Hunter (1984) and Mayer (1983).

None of the three supervisor and incumbent estimates of SD_y were statistically different from one another (Table 8). These results were similar to those reported by Mayer (1983), where Head Tellers (supervisors) provided estimates that were similar to Tellers' (incumbents) estimates in a large multibranch bank. Conversely, Matthews et al. (1990) found that supervisors provided estimates of SD_{y2} that were significantly larger than corresponding estimates provided by incumbents[2]. The findings of this study differ from previous research in that the two subject groups not only yielded similar percentile estimates, but similar estimates of SD_y as well. As noted, previous studies have reported mixed results.

Economic Utility

The second objective of this study was to estimate the utility of a testing process used to select applicants for the incumbent job class. Clearly, the results of the utility analysis support the position that test validity can translate into substantial performance/productivity savings to the organization. For the Eligibility Technician job class alone, these savings amount to over 6 million dollars (Table 11). The savings would be realized over the average tenure of an Eligibility Technician and the number of years the selection process was used. In this case, the savings would be realized over a 6.3 year period.

Top-Down Selection

A discussion of selection utility would not be complete without addressing the issues of top-down selection, and the ability of candidates to reject job offers (Boudreau, 1989; Hogarth & Einhorn, 1976; Murphy, 1986; Schmidt et al., 1984). Top-down selection has to do with making job offers based on

---

[2] Since SD_{y1} and SD_{y2} were found to be significantly different, Matthews et al. (1990) did not compute an estimate for SD_y, the primary estimate of interest in this study.

candidate rankings from selection process scores. To the degree that offers are not based on score ranking, and/or to the degree that candidates reject (or turn down) an offer of employment, selection utility is compromised. This occurs because an offer must then be (usually is) made to a candidate with a lower score ranking on the selection instrument, leading to lower job performance and fewer productivity savings. Remember, both scores on the selection process and supervisors judgements of incumbent job performance indicated that individuals with higher test scores not only perform better on the job, but they are more valuable to the organization).

For job classes where the selection ratio is high (i.e., the ratio of number hired to number applied is large), problems associated with top-down selection can lead to overestimates in utility by as much as 30 to 80 percent (Murphy, 1986). For the Eligibility Technician job class selection is made from the top-down based on candidates' score rankings. Some percentage of offers are rejected (roughly 1.2 interviews per hire), and some percentage of offers are made to candidates who possess bilingual skills, but who may have scored lower on the selection instrument than other candidates without the bilingual skills.

While more offers are made than accepted for the Eligibility Technician job class, top-down selection issues are minimal because of the large number of tied scores on the selection instrument. An average of 63 individuals (per test score) achieve the same score on the selection instrument. In short, this means that, in most cases, when an applicant rejects an offer of employment, another offer will be made to another applicant with the same predictor score. This does not compromise selection utility.

Conclusions and Recommendations

The findings of this study suggest that similar estimates of job worth can be obtained from multiple organizational levels. A possible explanation for this finding may stem from the fact that incumbents were chosen based on a minimum tenure of two years in the target job class. This limitation was imposed to include only those incumbents who would be eligible for promotion to a supervisory position based on tenure, suggesting that they would have a sufficient understanding of the job class. The results also substantiate the fact that utility analysis has the ability to provide management with valuable information concerning the costs and benefits of selection programs.

To better understand and explain the mixed results in the area of multi-level SD_y estimation, research should investigate the personal characteristics and job related dimensions used by individuals from different organizational levels in making estimates of worth, and how they weight them (Mathieu & Tannenbaum, 1989). This research should also address cognitive and decision-making processes as well. In addition, this research should take into consideration the level of complexity required to perform the target job class (Hunter, Schmidt, & Judiesch, 1990). Experimental designs could determine if job complexity and job dimension weights moderate SD_y estimates across organizational levels. There is already evidence to suggest that the standard deviation of employee output (SD_p) varies as a function of job complexity (Hunter et al., 1990). Research should also continue to compare various methods of SD_y estimation (Cascio & Ramos, 1986; Raju, Burke, & Normand, 1990) under the experimental design described.

References

Barrick, M. R., Barrett, G. V., Doverspike, D., Robison, S. J., & Grubs, L. L. (1990). Central tendency and its impact on three SDy procedures: A case study. *Journal of Occupational Psychology, 63*, 265-268.

Blankenship, M. H., Cesare, S. J., Giannetto, P. W., & Mandel, M. Z. (1991). *An investigation of SDy estimates: Equivalence across methods*. Paper presented at the annual meeting of the American Psychological Association, San Francisco, CA.

Bobko, P., Karren, R., & Kerkar, S. P. (1987). Systematic research needs for understanding supervisory-based estimates of SD_y in utility analysis. *Organizational Behavior and Human Decision Processes, 40*, 69-95.

Bobko, P., Karren, R., & Parkington, J. J. (1983). Estimation of standard deviations in utility analyses: An empirical test. *Journal of Applied Psychology, 68*, 170-176.

Boudreau, J. W. (1989). Selection utility analysis: A review and agenda for future research. In M. Smith & I. T. Robertson (Eds.), *Advances in selection and assessment* (pp. 227-257). New York: John Wiley & Sons.

Brogden, H. E. (1949). When testing pays off. *Personnel Psychology, 2*, 171-183.

Burke, M. J., & Frederick, J. T. (1984). Two modified procedures for estimating standard deviations in utility analyses. *Journal of Applied Psychology, 69*, 482-489.

Burke, M. J., & Frederick, J. T. (1986). A comparison of economic utility estimates for alternate SD_y estimation procedures. *Journal of Applied Psychology, 71*, 334-339.

Cascio, W. F. (1987). *Costing human resources: The financial impact of behavior in organizations* (rev. ed.). Boston: Kent.

Cascio, W. F., & Ramos, R. A. (1986). Development and application of a new method for assessing job performance in behavioral/economic terms. *Journal of Applied Psychology, 71*, 20-28.

Cascio, W. F., & Silbey, V. (1979). Utility of the assessment center as a selection device. *Journal of Applied Psychology, 64*, 107-118.

Cesare, S. J., Blankenship, M. H., Mandel, M. Z., & Giannetto, P. W., (1991). A predictive validation study of the methods used to select eligibility technicians. *A psychometric and fiscal assessment for selecting eligibility workers: Multiple agency results*. Symposium conducted at the annual meeting of the International Personnel Management Association Assessment Council, Chicago, Illinois.

Cronbach, L. J., & Gleser, G. C. (1965). *Psychological tests and personnel decisions* (2nd ed.). Urbana: University of Illinois Press.

Cronshaw, S. F., & Alexander, R. A. (1985). One answer to the demand for accountability: Selection utility as an investment decision. *Organizational Behavior and Human Decision Processes, 35*, 102-118.

DeSimone, R. L., Alexander, R. A., & Cronshaw, S. F. (1986). Accuracy and reliability of SD_y estimates in utility analysis. *Journal of Occupational Psychology, 59*, 93-102.

Edwards, J. E., Frederick, J. T., & Burke, M. J. (1988). Efficacy of modified CREPID SD_ys on the basis of archival organizational data. *Journal of Applied Psychology, 73,* 529-535.

Greer, O. L., & Cascio, W. F. (1987). Is cost accounting the answer? Comparison of two behaviorally based methods for estimating the standard deviation of job performance in dollars with a cost-accounting-based approach. *Journal of Applied Psychology, 72,* 588-595.

Hayes, W. L. (1963). *Statistics for psychologists*. New York: Holt, Rinehart, & Winston.

Hogarth, R. M., & Einhorn, H. J. (1976). Optimal strategies for personnel selection when candidates can reject job offers. *Journal of Business, 49,* 479-495.

Hunter, J. E., & Hunter, R. F. (1984). Validity and utility of alternate predictors of job performance. *Psychological Bulletin, 96,* 72-98.

Hunter, J. E., Schmidt, F. L., & Judiesch, M. K. (1990). Individual differences in output as a function of job complexity. *Journal of Applied Psychology, 75,* 28-42.

Landy, F. J., Farr, J. L., & Jacobs, R. R. (1982). Utility concepts in performance measurement. *Organizational Behavior and Human Performance, 30,* 15-40.

Mathieu, J. E., & Tannenbaum, S. I. (1989). A process-tracing approach toward understanding supervisors' SDy estimates: Results from five job classes. *Journal of Occupational Psychology, 62,* 249-256.

Matthews, M. D., Looper, L. T., & Engquist, S. K. (1990). *Applicability of utility models to the evaluation of military manpower and personnel research programs: A critical review and illustrations* (AFHRL-TP-90-54). Air Force Systems Command: Brooks Air Force Base, Texas.

Mayer, R. S. (1983). *An evaluation of alternative methods of estimating the standard deviation of job performance to determine the utility of a test in a fixed-treatment sequential employee selection process.* Paper presented at the annual meeting of the International Personnel Management Association Assessment Council, Washington, D. C.

Murphy, K. R. (1986). When your top choice turns you down: Effect of rejected offers on the utility of selection tests. *Psychological Bulletin, 99,* 133-138.

Raju, N. S., Burke, M. J., & Normand, J. (1990). A new approach for utility analysis. *Journal of Applied Psychology, 75,* 3-12.

Reilly, R. R., & Smither, J. W. (1985). An examination of two alternate techniques to estimate the standard deviation of job performance in dollars. *Journal of Applied Psychology, 70,* 651-661.

Roche, W. J. (1961). *The Cronbach-Gleser utility function in fixed-treatment employee selection*. Unpublished doctoral dissertation, Southern Illinois University, Carbondale, IL.

Schmidt, F. L., Hunter, J. E., McKenzie, R. C., & Muldrow, T. W. (1979). Impact of valid selection procedures on work-force productivity. *Journal of Applied Psychology, 64,* 609-626.

Schmidt, F. L., Hunter, J. E., & Pearlman, K. (1982). Assessing the economic impact of personnel programs on workforce productivity. *Personnel Psychology, 35,* 333-347.

Schmidt, F. L., Mack, M. J., & Hunter, J. E. (1984). Selection utility in the occupation of U.S. park ranger for three modes of test use. *Journal of Applied Psychology, 69*, 490-497.

Vance, R. J., & Colella, A. (1990). The utility of utility analysis. *Human Performance*, 3, 123-139.

Weekley, J. A., Frank, B., O'Connor, E. J., & Peters, L. H. (1985). A comparison of three methods of estimating the standard deviation of performance in dollars. *Journal of Applied Psychology, 70*, 122-126.

Zeidner, J., & Johnson, C. D. (July, 1989). *The utility of selection for military and civilian jobs*. Institute for Defense Analyses (P-2239): Alexandria, Virginia.